Available online at www.sciencedirect.com
SCIENCE DIRECT
Molecular Phylogenetics and Evolution 40 (2006) 459–470
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Phylogenetic utility of indels within ribosomal DNA and β-tubulin sequences from fungi in the *Rhizoctonia solani* species complex

Dolores González [a,*], Marc A. Cubeta [b], Rytas Vilgalys [c]

[a] *Departamento de Biodiversidad y Sistemática, Instituto de Ecología, A.C., Apartado Postal 63, Xalapa Veracruz 91000, México*
[b] *Department of Plant Pathology, North Carolina State University, Raleigh, NC 27695-7567, USA*
[c] *Department of Biology, Duke University, Durham, NC 27708-0338, USA*

Received 15 July 2005; revised 16 February 2006; accepted 16 March 2006
Available online 2 May 2006

Abstract

The genus *Rhizoctonia* consists of a diverse assemblage of anamorphic fungi frequently associated with plants and soil throughout the world. Some anamorphs are related with teleomorphs (sexual stage) in different taxonomic classes, orders, and families. The fungus may exist as pathogen, saprophyte, or mycorrhizal symbiont and shows extensive variation in characteristics such as geographic location, morphology, host specificity, and pathogenicity. In this study, phylogenetic analyses were performed in the *Rhizoctonia solani* species complex with individual and combined data sets from three gene partitions (ITS, LSU rDNA, and β-tubulin). To explore whether indels were a source of phylogenetically informative characters, single-site indels were treated as a new state, while indels greater than one contiguous nucleotide were analyzed by including them as ambiguous data (Coding A); excluding them from the analyses (Coding B), and with three distinct codes: multistate for different sequence (Coding C); multistate for different length (Coding D) and different characters for each distinct sequence (Coding E). Results suggest that indels in noncoding regions contain phylogenetic information and support the fact that the *R. solani* species complex is not monophyletic. Six clades within *R. solani* (teleomorph = *Thanatephorus*) representing distinct anastomosis groups and five clades within binucleate *Rhizoctonia* (teleomorph = *Ceratobasidium*) were well supported in all analyses. The data suggest that clades with representatives of *R. solani* fungi belonging to anastomosis groups 1, 4, 6, and 8 should be recognized as phylogenetic species.

Keywords: Alignment; Sequence homology; Parsimony; Coding indels

1. Introduction

The genus *Rhizoctonia* include a wide variety of genetically diverse microscopic fungi associated with plant roots and soil that are: (1) economically important plant pathogens, (2) saprophytes on organic matter, and (3) symbionts associated with orchids, mosses and liverworts (Cubeta and Vilgalys, 2000; Stalpers and Andersen, 1996). *Rhizoctonia solani* is the most recognized species. It causes diseases on more than 200 species of plants including rice, corn, wheat, soybeans, potato, cotton, fruit and forest trees, turfgrasses, and ornamentals and shows extensive variation in characteristics such as geographic location, morphology, host specificity, and pathogenicity. The extent of this variation has been considered as evidence that this fungus is a species complex. However, uncertainty exists about how many species are there within this complex. Historically, fungi in the *R. solani* complex have been characterized and identified based on the concept of "hyphal anastomosis" which was developed more than 70 years ago (Matsumoto et al., 1932; Schultz, 1936). The concept implies that isolates that have the ability to undergo hyphal fusion (i.e., anastomose) are genetically related. Hyphal anastomosis represents a type of somatic recognition in *Rhizoctonia* that is analogous to vegetative (somatic) incompatibility systems previously described in other filamentous fungi. The criterion of hyphal anastomosis has been used extensively to place isolates of *R. solani* (teleomorph = *Thanatephorus*) and binucleate *Rhizoc-*

* Corresponding author. Fax: +52 228 818 7809.
E-mail address: dolores@ecologia.edu.mx (D. González).

1055-7903/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2006.03.022

tonia (teleomorph = *Ceratobasidium*) into distinct groups called anastomosis groups (AG). Recent protein and DNA sequence-based studies support the separation of *R. solani* and binucleate *Rhizoctonia* into genetically distinct groupings, but have also revealed considerable genetic diversity within an AG (González et al., 2001; Kuninaga et al., 1997).

DNA sequence-based studies have provided insight into phylogenetic relationships in a wide variety of organisms (i.e., Hibbett and Thorn, 2001; Milinkovitch et al., 1994; Moncalvo et al., 2002; Rehner and Buckley, 2005; Soltis et al., 2000). However, alignment of sequences is still a matter of concern in phylogenetic analyses especially when dealing with sequences of different length. In this case, alignment of homologous sequences is problematic because gaps are introduced to the data matrix. In most phylogenetic analyses gaps are ignored as ambiguous characters or are simply deleted from the data sets arguing that they are phylogenetically uninformative (Olsen, 1988; Olsen and Woese, 1993). But in studies where indels are considered as putative insertion or deletion events, they are incorporated in the analyses as a class of phylogenetic characters (i.e., Aagesen et al., 2005; Grubisha et al., 2002; Hibbett et al., 1995; Lutzoni et al., 2000; Simmons and Ochoterena, 2000). Some authors include multibase indels as ambiguous characters and then repeat the analysis coding them as binary characters (i.e., Kropp et al., 1997; Milinkovitch et al., 1994). Others include them as fifth state increasing the relative weight of gaps according to their length (Coetzee et al., 2003). Very few however, explore levels of variation among indels and use a coding scheme to represent such variation (i.e., González, 1996; Hibbett et al., 1995; Simmons and Ochoterena, 2000; Vogler and DeSalle, 1994).

In a previous study to investigate the phylogenetic relationships within *R. solani* (teleomorph = *Thanatephorus*) and binucleate *Rhizoctonia* spp. (teleomorph = *Ceratobasidium*), we found that sequence alignment of the ITS and LSU regions of the ribosomal DNA generated several gaps of different length (González et al., 2001). The decision for choosing among possible alignments, especially in the ITS region, was problematic. Therefore a conservative approach was employed for parsimony analyses and for selecting the alignment with the highest penalty for introducing gaps and for the length of the gaps. The indels generated were coded as ambiguous characters. Under these parameters our conclusions were that *R. solani* was not monophyletic and that four unrelated clades represented well defined genetically isolated groups (González et al., 2001).

In this study, we tested the potential utility of gaps as phylogenetic characters by assessing their effect on the stability of phylogenetic hypothesis using an expanded data set with more taxa to our previous data matrix (González et al., 2001). To explore if indels are a source of useful phylogenetic characters, single-site indels were treated as a new state, while indels greater than one contiguous nucleotide were analyzed by including them as ambiguous data (Coding A); excluding them from the analyses (Coding B); and with three distinct codes: multistate for different sequence (Coding C); multistate for different length (Coding D); and different characters for each distinct sequence (Coding E).

2. Materials and methods

2.1. Taxon sampling and extraction, amplification, and sequencing of DNA

Forty-three isolates representing 28 anastomosis groups (AG) of binucleate *Rhizoctonia* spp., (teleomorph = *Ceratobasidium*) and multinucleate *R. solani* (teleomorph = *Thanatephorus*) were used in this study. The sequence of the ITS and LSU rDNA regions were previously determined for 37 of these isolates (González et al., 2001). The ITS and LSU rDNA regions were sequenced for 6 additional isolates of *Rhizoctonia*, and a section of the β-tubulin gene of 368 nucleotides was sequenced for all 43 isolates. Isolates for this study were obtained from USA, Japan, Canada, Australia and Scotland (Table 1).

DNA extraction, amplification and sequencing were performed as described in González et al. (2001). Primers named "B36F" (5′-CACCCACTCCCTCGGTGGTG-3′), and "B12R" (5′-CATGAAGAAGTGAAGACGCGGGA A-3′) were used to amplify and/or sequence the β-tubulin region for 60% of isolates. Two additional primers "BTUB-1F" (5′-CACCCACTCWCTWGGTGGT-3′) and "BTUB-380R" (5′-TACCCATGTTGACAGCRAG-3′) were designed and used to amplify isolates when difficulties were encountered in obtaining an amplified PCR product with the B36F and B12R primers.

2.2. Alignment of DNA sequence data

Sequences were aligned using the Clustal V program (Higgins et al., 1992) within the Megalign software package (Lasergene, DNASTAR Inc). The aligned data set consisted of 43 OTUs and 2049 nucleotide positions from the ITS, LSU, and β-tubulin genes. A survey of primary homology assessment was performed on 12 alignments with combinations of penalties of 10, 20, 50, and 100 for introducing gaps and penalties of 5, 10, 20, 50, and 100 for the length of gaps. Phylogenetic analyses were performed initially by coding indels of one nucleotide as an additional character state and recoding indels of two or more contiguous nucleotides. In these analyses, a "dash" symbol (-) was added in the "matrix format" in PAUP* ver. 4.0b4 (Swofford, 2000) in order to consider indels of one nucleotide as a different character state. For indels of two or more contiguous nucleotides, positions in the matrix that contained gapped sequences were identified, and all gaps from these positions were removed and recoded. After removing gaps, some taxa have one combination of nucleotides, others have a different combination, and others lack that sequence altogether. The indels were analyzed by including them as ambiguous data (Coding A) and excluded them from the analyses (Coding B). Then, three approaches were used to describe the variation of the

Table 1
Anastomosis group (AG/CAG) and subgroup designation, origin, and source of isolates of *Rhizoctonia* species with *Ceratobasidium* and *Thanatephorus* teleomorphs used for current sequence analysis of the internal transcribed spacer (ITS1 and ITS2), 28S large subunit (LSU) ribosomal DNA (rDNA) regions and a section of the β-tubulin gene together with GenBank accession number

Anastomosis group/subgroup	Isolate	Origin (source)[a]	GenBank Accession Nos.	
			rDNA (ITS and 28S)	β-Tubulin
Rhizoctonia solani (teleomorph = Thanatephorus)				
AG-1-IA	1Rs (ATCC 66159)	Soybean, US (11)	AF354060	DQ085463
AG-1-IB	36Rs (ATCC 66150)	Turfgrass, US (2)	AF354059	DQ085464
AG-1-IC	3Rs (ATCC 44661)	Pine, Canada (1)	AF354058	DQ085465
AG-2-1	8Rs (ATCC 44658)	Soil, Australia (1)	AF354063	DQ085466
AG-2-2	9Rs (ATCC 44659)	Carrot, US (1)	DQ097887	DQ085467
AG-2-2 IV	16Rs	Sugar beet, Japan (10)	AF354117	DQ085468
AG-3	4Rs (ATCC 14006)	Potato, US (9)	AF354064	DQ085469
AG-4 HGI	AH-1 (ATCC 76126)	Peanut, Japan (7)	AB000012, AF354118	DQ085488
AG-4 HGII	7Rs (ATCC 44662)	Alfalfa, US (1)	AF354074	DQ085472
AG-4 HGII	18Rs	Sugar beet, Japan (10)	AF354072	DQ085470
AG-4 HGII	30Rs (ATCC 48803)	Unknown, Canada (3)	AF354073	DQ085471
AG-4 HGIII	6Rs (ATCC 42127)	Conifer, US (3)	AF354077	DQ085473
AG-4 HGIII	45Rs (ATCC 10177)	Sugar beet, US (6)	AF354076	DQ085474
AG-5	10Rs	Soybean, Japan (10)	AF354078	DQ085475
AG-6 HG-I	72Rs	Soil, Japan (10)	AF354061	DQ085476
AG-6 GV	74Rs	Soil, Japan (10)	AF354062	DQ085477
AG-7	76Rs	Soil, Japan (4)	AF354096	DQ085478
AG-8	(ZG1-2)SA50	Oats, Australia (5)	AF354067	DQ085479
AG-8	(ZG1-3)SA1512	Barley, Australia (12)	AF354068	DQ085480
AG-8	(ZG1-5)92547	Barley, Australia (12)	DQ097888	DQ085481
AG-9	116Rs	Potato, US (4)	AF354065	DQ085482
AG-10	W45b3	Wheat, US (10)	AF354111	DQ085484
AG-10	(ZG9)91614	Barley, Australia (8)	AF354071	DQ085483
AG-11	(ZG-3)R1013	Lupine, Australia (13)	AF354079	DQ085485
AG-BI	22Rs	Soil, Japan (10)	AF354070	DQ085486
AG-BI	TE 2-4	Soil, Japan (10)	AB000044	DQ085487
Binucleate Rhizoctonia spp. (teleomorph = Ceratobasidium)				
AG-A	C-662	Soil, Japan (10)	AF354092	DQ085489
AG-Ba	C-460	Rice, Japan (10)	AF354088	DQ085490
AG-Bb	C-455	Rice, Japan (10)	AF354087	DQ085491
AG-Bo	SIR-2	Sweetpotato, Japan (10)	AF354091	DQ085492
AG-D	C-610	Unknown, Japan (10)	AF354090	DQ085493
AG-F	SIR-1	Sweetpotato, Japan (10)	AF354085	DQ085494
AG-G	C-653	Unknown, Japan (10)	DQ097889	DQ085495
AG-H	STC-9	Soil, Japan (10)	AF354089	DQ085496
AG-L	FK02-1	Soil, Japan (10)	AF354093	DQ085497
AG-O	FK06-2	Soil, Japan (10)	AF354094	DQ085498
AG-Q	C-620	Soil, Japan (10)	AF354095	DQ085499
CAG-1	BN1	Turfgrass, US (2)	AF354086	DQ085500
CAG-3	BN31	Peanut, US (2)	AF354080	DQ085501
CAG-4	BN38	Soybean, US (2)	AF354081	DQ085502
CAG-5	BN37	Cucumber, US (2)	AF354082	DQ085503
CAG-6	BN74 (ATCC 13247)	*Erigeron*, US (2)	AF354083	DQ085504
CAG-7	BN22 (FL FTCC585)	*Pittosporum*, US (2)	AF354084	DQ085505

[a] Isolates provided by: 1, N. Anderson; 2, L. Burpee; 3, E. Butler; 4, D. Carling; 5, A. Dube; 6, J. Kotila; 7, S. Kuninaga; 8, G. MacNish; 9, G. Papavizas; 10, A. Ogoshi; 11, N. O'Neill; 12, S. Neate; and 13, M. Sweetingham.

indels. First, a single multistate treatment as described by Lutzoni et al. (2000) was used (Coding C). Second, a multistate code to identify only a specific length of the indel (Coding D); and third a different character for describing a specific nucleotide sequence (Coding E, Fig. 1). With these coding approaches, the origin of indels of two or more positions could be interpreted as a single insertion–deletion event. Seventeen analyses were performed with different combination of data sets. Three analyses corresponded to each gene partition, three to each possible combination of two gene partitions and one for the three data sets combined. Each of these analyses were carried out with indels as ambiguous characters (Coding A) and with indels excluded from the analyses (Coding B). Codings C, D, and E were applied only in the combined data set (Table 3). Indels that presented one or more unidentified nucleotides (N) in the sequence (regardless of sequence length) were coded as an ambiguous character.

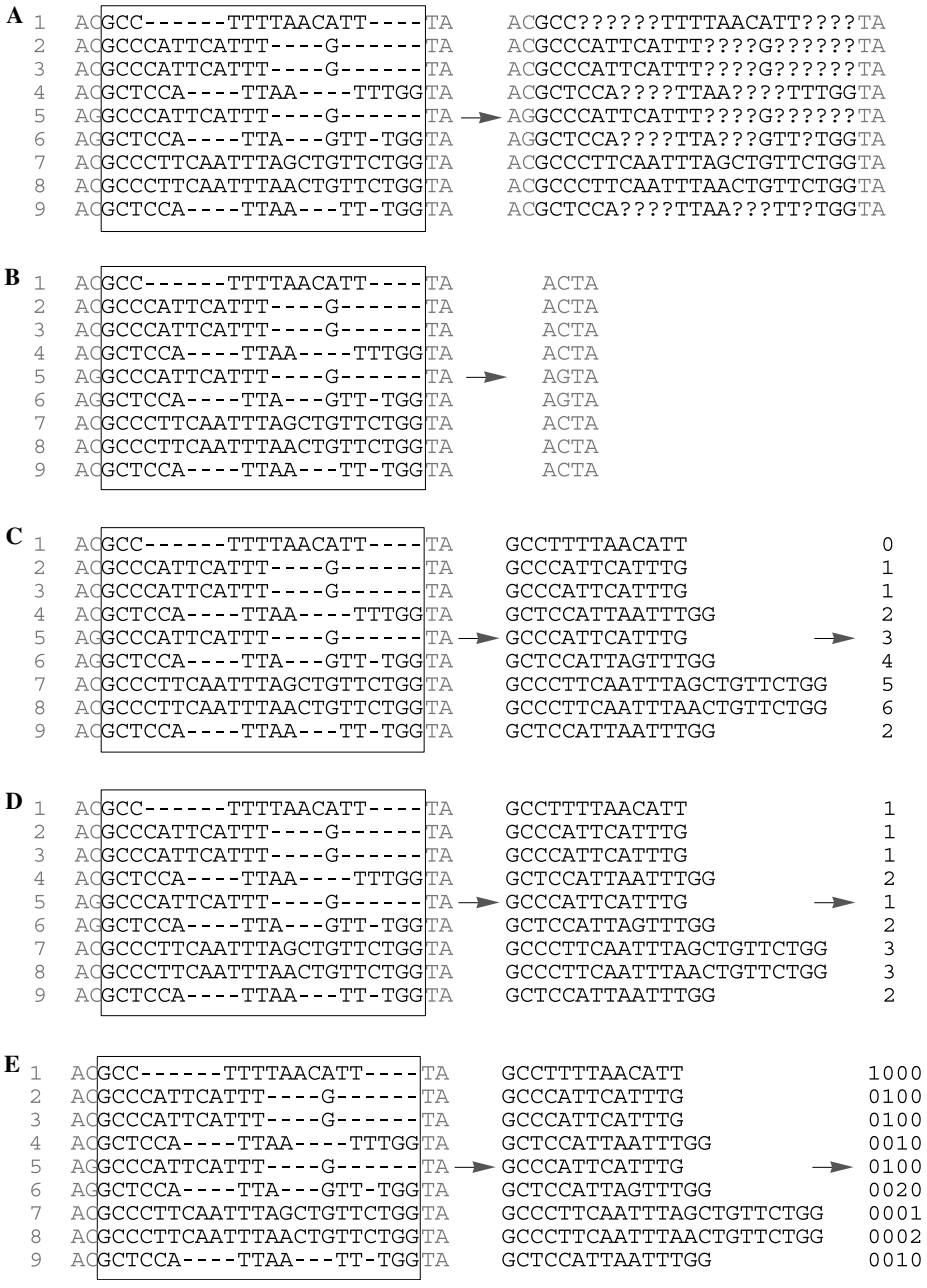

Fig. 1. Different codes used for indels of two or more contiguous nucleotides. Positions in the data matrix that contain gapped sequences were identified. These positions were removed and indels were recoded as follows: (A) as ambiguous characters ("?"); (B) excluded; (C) as multistate for different sequence; (D) as multistate for different length; and (E) as different characters for each distinct sequence.

2.3. Phylogenetic analyses

All analyses were performed using the maximum parsimony criterion in PAUP* ver. 4.0b4 (Swofford, 2000). We used heuristic searches with 500 random addition replicates and TBR branch swapping after exclusion of uninformative characters. MulTrees option was in effect and zero length branches collapsed. MAXTREES was set to autoincrease in all analyses. Branch support was assessed by Bremer support index, and bootstrap analysis based on 1000 replicates and 10 heuristic searches per replicate. MAXTREES was set to 10000 in bootstrap analyses. Despite the limitations of the bootstrap, it provides an indication of internal support (Sanderson, 1989). Data decisiveness (DD) (Kitching et al., 1998) was calculated as a measure of information content in individual and combined data sets.

Because appropriate outgroups were not available, preliminary higher and lower-level cladistic analyses were performed with LSU rDNA sequences (≈1000 nucleotides) from GenBank. Higher-level analysis included 95 representatives of the euagaric clade in the homobasidiomycetes (Hibbett and Thorn, 2001), and 8 sequences from *Thanatephorus* and *Ceratobasidium* from this study. This analysis

was intended to resolve the placement of *Thanatephorus* and *Ceratobasidium* among 82 homobasidiomycete genera. In the preliminary higher-level analysis, the species of *Thanatephorus* and *Ceratobasidium* used were nested with members of the Stereales, Aphylophorales, and Cantharellales. Once we found a possible sister relation, lower-level analysis was performed with ≈600 nucleotides of the LSU. This analysis included all 43 sequences from *Thanatephorus* and *Ceratobasidium* used in this study and 58 additional sequences from GenBank; 16 of these sequences corresponded to Ceratobasidiales (*Uthatobasidium*, *Thanatephorus*, and *Ceratobasidium*); 14 to Stereales (*Botryobasidium*); 14 to the Aphylophorales (*Sistotrema*) and 14 to the Cantharellales (*Clavulina* and *Multiclavula*). Based on the results obtained in our lower-level analyses with the LSU rDNA sequences, the orientation of the trees was performed with a binucleate *Rhizoctonia* sp. (AG-H) (teleomorph = *Ceratobasidium*).

3. Results

3.1. Sequence analyses

Among the 12 alignments generated, we chose those that have identical 5′ and 3′ ends of the 5.8S region of the rDNA. From these, the alignment with a gap penalty of 20 and a gap length penalty of 10 was selected. This alignment generated 74 indels of one nucleotide and 34 indels of two or more contiguous nucleotides (Table 2). After alignment, the combined ITS + LSU + β-tubulin data set had 2049 characters. From these 2049 characters, 1398 (68%) were constant, 196 (10%) were variable but parsimony uninformative, and 455 (22%) were parsimony informative (Coding A). The removal of indels (Coding B) reduced informative characters to 296 (14%). Recoding indels as multistate characters (Coding C) produced 2083 characters, 368 (18%) were informative; while using a multistate code to identify an indel of specific length (Coding D) originated 367 informative characters. Indels coded as different characters for describing a specific nucleotide sequence (Coding E) resulted in 2301 characters. From these 476 (21%) were informative. The data decisiveness (DD) score was higher for the β-tubulin sequences than any other gene partition. The lowest DD score corresponded to the three gene partitions combined and all indels excluded (Table 3).

Nucleotide composition for the combined data matrix with all characters included (2049 nucleotides), varied little among taxa ($A = 0.256$, $C =$;0.222, $G = 0.247$, and $T = 0.273$), with an average $A + T$ composition of 53%. The β-tubulin gene region was the gene partition that had the largest nucleotide composition variation ($A = 0.058$, $C = 0.481$, $G = 0.178$, and $T = 0.281$), with an average $A + T$ composition of 34%. The nucleotide composition was homogeneous based on a χ^2 test of homogeneity of base frequencies across taxa. The value obtained for the combined data matrix and all characters included was 22.993 (df = 126, $P = 1.0000$). Homogeneity of the frequency of nucleotides was also seen for each gene partition and for the combined data matrix with indels excluded.

3.2. Phylogenetic analyses

Analyses were executed using only the binucleate *Rhizoctonia* sp. (AG-H) (teleomorph = *Ceratobasidium*) as outgroup. The selection of this isolate was based on the results obtained in our higher and lower-level analyses (data not shown) with LSU rDNA sequences, where isolate AG-H resulted the sister group to the rest of the isolates included in this study. A comparison of general features of the most parsimonious trees (MPTs) found in all 17 analyses performed is summarized in Table 3. Different levels of resolution for phylogenetic relationships were reconstructed in our analyses with the three data sets taken individually and when indels were included as ambiguous characters or excluded. The strict consensus tree obtained with the LSU sequences and indels excluded had the poorest resolution due to 11,323 competing topologies. Even the consensus of 113,776 trees found with the 368 nucleotides sequenced from the β-tubulin gene had greater resolution in basal branches than the LSU tree. In contrast, the strict consensus of the ITS sequences was well resolved and supported in terminal branches but less resolved in basal branches. In the analyses of each gene partition there was better resolution when the nucleotide positions with indels were included. Similar results were obtained for the combined analysis of LSU and β-tubulin sequences. In contrast, other combinations of two gene partitions (ITS + LSU and ITS + TUB) had better resolution in basal branches when these positions were excluded but they were not supported by bootstrap analyses.

From all 17 analyses (individual or combined data sets with indels included as ambiguous character, excluded, or recoded), the greatest resolution was obtained when the three data partitions were combined, and gapped positions

Table 2
Number of positions in the sequence of the internal transcribed spacer (ITS1 and ITS2), 28S large subunit (LSU) ribosomal DNA (rDNA) regions and a section of the β-tubulin gene before and after alignment with the Clustal V program (Higgins et al., 1992)

Gene partition	Before		After	Indels of one nucleotide	Indels of two or more contiguous nucleotides
	Shortest sequence	Longest sequence			
ITS	571	654	730	21	23
LSU	920	936	954	53	10
β-Tubulin	313	365	365	0	1

Table 3
Results of seventeen analyses of the internal transcribed spacer (ITS), 28S large subunit (LSU), and β-tubulin under different combinations of data sets and indel coding schemes (A, B, C, D, E)

Gene partition	# Of PIS	# Of MPT	Length of MPT	CI	RI	DD
ITS						
A	289	7	1255	0.4191	0.6240	0.5691
B	144	29	491	0.4603	0.6764	0.6394
LSU						
A	93	4726	266	0.4737	0.6585	0.6153
B	79	11,323	221	0.4842	0.6535	0.6135
β-Tubulin						
A	73	>113,776	262	0.4084	0.7002	0.6509
B	73	>113,776	262	0.4084	0.7002	0.6509
ITS + LSU						
A	382	8	1568	0.4158	0.6100	0.5544
B	223	86	754	0.4416	0.6333	0.5916
ITS + β-tubulin						
A	362	78	1706	0.3710	0.5631	0.4976
B	217	114	910	0.3659	0.5681	0.5110
LSU + β-tubulin						
A	166	28	641	0.3635	0.5599	0.4950
B	152	2660	580	0.3690	0.5674	0.5054
ITS + LSU + β-tubulin						
A	455	6	2028	0.3743	0.5572	0.4924
B	296	177	1190	0.3697	0.5495	0.4918
C	368	24	1674	0.4337	0.5734	0.5187
D	367	15	1590	0.3975	0.5653	0.5097
E	476	9	1739	0.3755	0.5748	0.5184

We report the number of phylogenetically informative sites (PIS); most parsimonious trees (MPT); length of the MPT, consistency index (CI); retention index (RI); and data decisiveness (DD) score.

were included and coded as ambiguous characters (Coding A, 6 trees) or when indels were recoded. From these last analyses, the greatest resolution was when indels were coded using a different character for describing the combination of nucleotide of the sequence (Coding E, 9 trees). In contrast, the analysis with three gene partitions but gapped positions excluded (Coding B) recovered a strict consensus with 177 trees (Fig. 2).

All phylogenetic analyses revealed numerous well-supported terminal groupings that correspond with previously recognized AG or AG subgroups within *Thanatephorus* and *Ceratobasidium* with *Rhizoctonia* anamorphs (Figs. 2 and 3). The clades obtained with bootstrap values above 50% in at least five of the 12 analyses for each gene partition and in combination when positions with gaps were included or excluded are described in Table 4. Six clades of isolates of the same AG of *R. solani* (teleomorph = *Thanatephorus*) (AG-1, AG-4, AG-6, AG-8, AG-10, and AG-BI), one clade of subgroup AG-2-2 and five clades with two or three binucleate *Rhizoctonia* spp. (teleomorph = *Ceratobasidium*) (AG-D, CAG-1; AG-Q, AG-Ba, AG-Bb; AG-G, AG-L, AG-O; AG-A, AG-Bo, and AG-F, CAG-5) were always recovered with high bootstrap and Bremer support values. In clades containing isolates of AG-1, AG-4, AG-8, and clades including AG-Q AG-Ba AG-Bb, and AG-G AG-L AG-O internal relationships of AG or subgroups varied.

In each consensus tree obtained when the three gene partitions are combined, six isolates (CAG-3, AG-F, CAG-5, CAG-6, CAG-4, and CAG-7) from binucleate *Rhizoctonia* (teleomorph = *Ceratobasidium*) are mixed and dispersed across several small clades of *R. solani* (teleomorph = *Thanatephorus*), indicating these groups are not monophyletic. In all analyses, trees always showed several small clades containing only isolates of *Ceratobasidium* and one large clade with high bootstrap and Bremer support indices containing all 26 isolates of *R. solani* (AG-1 to 11 and AG-BI) and the six isolates of *Ceratobasidium* (CAG-3, CAG-4, CAG-5, CAG-6, CAG-7, and AG-F). This large clade contains eight well-supported small clades (Table 5). One clade included only *R. solani* (AG-2-2). A second clade incorporated two binucleate *Ceratobasidium* isolates (AG-F and CAG-5). The six remaining clades included only isolates of a single anastomosis group of *R. solani* (AG-1, AG-4, AG-6, AG-8, AG-10, and AG-BI). These eight clades are also present in most analyses with one or two gene partitions (Table 4). This indicates that there are phylogenetically informative regions in the three gene partitions used and such clades are relatively stable to different strategies for coding the indels. Despite the constancy of the eight clades, relationships among them are not stable. Most variation among trees was due to local rearrangement of AG (Figs. 2 and 3). It is clear that the placement of clades is sensitive to gene partition and the strategy used for coding indels.

4. Discussion

The great variety of fungi in the *R. solani* species complex has caused problems in delimiting species within this

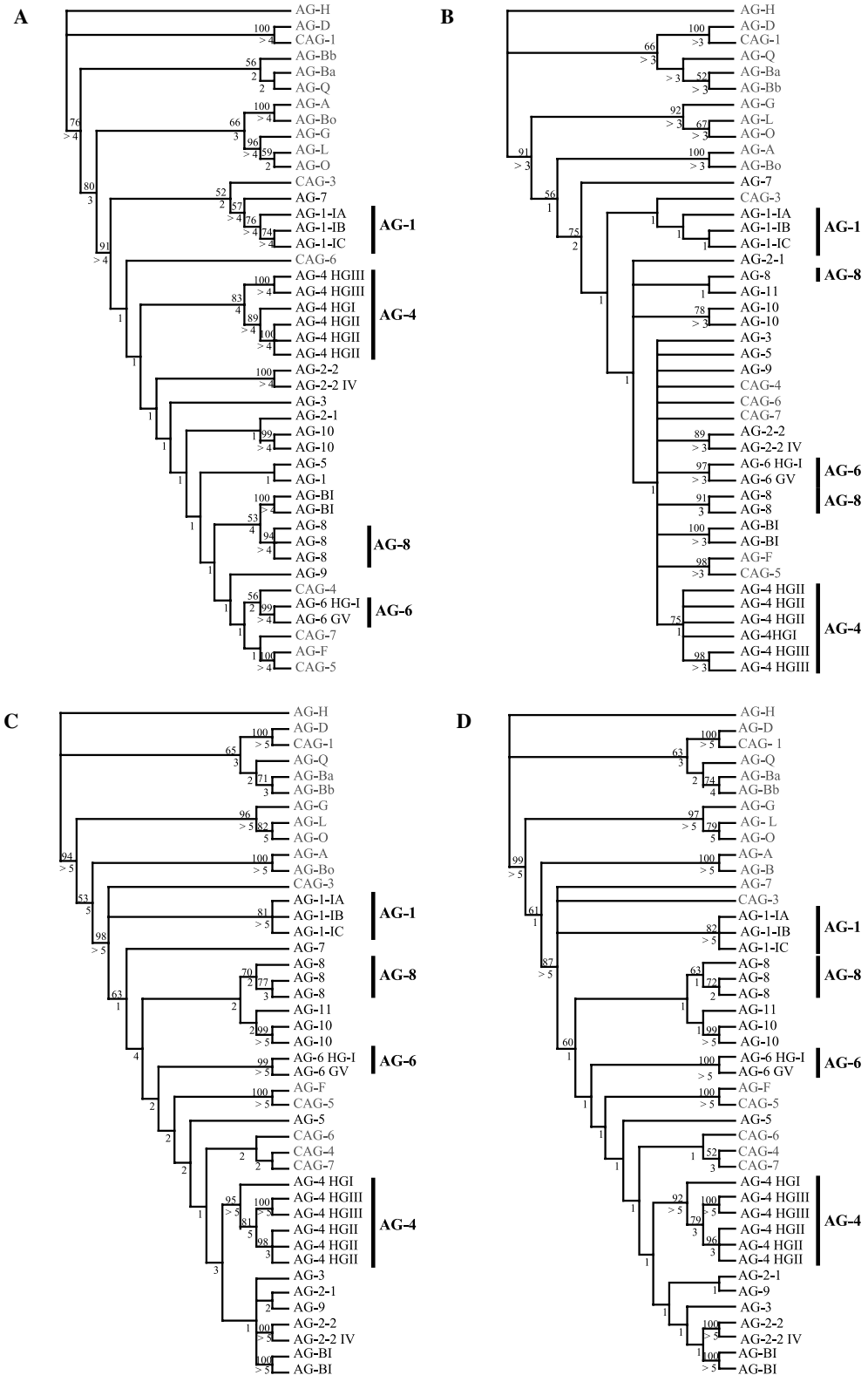

Fig. 2. Strict consensus trees resulting from combined analyses of internal transcribed spacer, 28S large subunit, and β-tubulin gene partitions including indels with coding schemes (A–D). In black, isolates of *Rhizoctonia solani* (teleomorph = *Thanatephorus*); in gray, isolates of binucleate *Rhizoctonia* spp. (teleomorph = *Ceratobasidium*).

species-rich complex. Cladistic analyses of individual and combined data sets from three gene partitions (ITS, LSU rDNA, and β-tubulin) confirmed that *R. solani* (teleomorph = *Thanatephorus*) and binucleate *Rhizoctonia* (teleomorph = *Ceratobasidium*) are not monophyletic groups (González et al., 2001). Instead, a clade containing

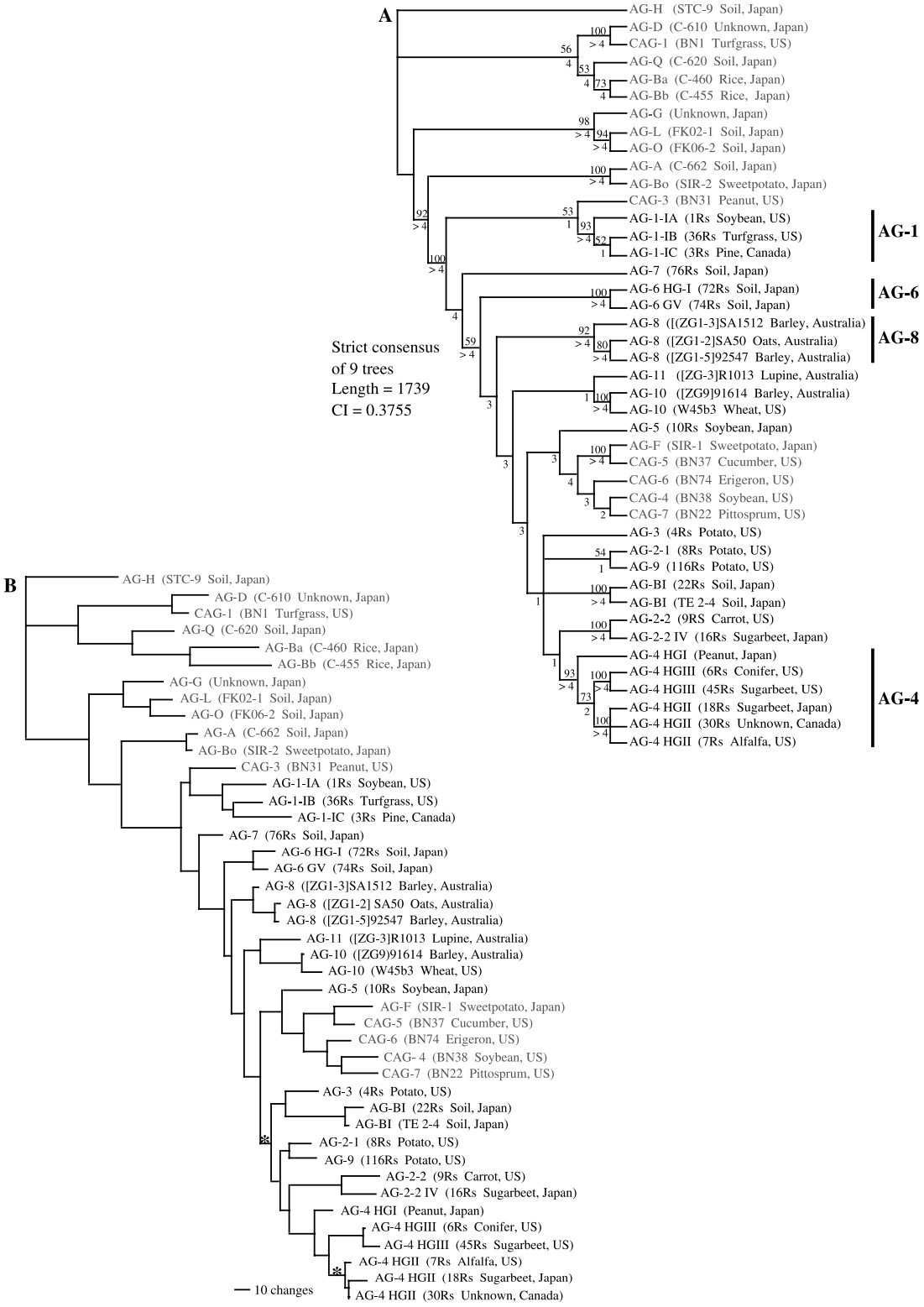

Fig. 3. (A) Strict consensus tree resulting from combined analyses of internal transcribed spacer, 28S large subunit, and β-tubulin gene partition including indels recoded as different characters for each distinct sequence (Coding E). (B) One of 9 equally parsimonious trees (Length = 1739, CI = 0.3755, RI = 0.5748). Nodes that collapsed in the strict consensus tree are marked with an asterisk above the branch. In black, isolates of *Rhizoctonia solani* (teleomorph = *Thanatephorus*); in gray, isolates of binucleate *Rhizoctonia* spp. (teleomorph = *Ceratobasidium*).

all 26 isolates of *R. solani* (AG-1 to 11 and BI) and six binucleate (CAG-3, CAG-4, CAG-5, CAG-6, CAG-7, and AG-F) was always recovered with high bootstrap and Bremer support indices. These results suggest that at least some isolates of *Ceratobasidium* may be more closely related with other isolates from *Thanatephorus*. This close relationship

Table 4
Effect of indels in clade support from partitioned analyses

CLADE	A						B					
	ITS	LSU	TUB	ITS + LSU	ITS + TUB	LSU + TUB	ITS	LSU	TUB	ITS + LSU	ITS + TUB	LSU + TUB
AG-A, AG-Bo	100	100		100	100	100	100	100		100	100	100
AG-G, AG-L, AG-O	84	93		99	71	82	98	82		98	78	54
AG-L, AG-O		98		97			69	99		99		82
AG-G, AG-L,			58		84	88			58		69	
AG-D, CAG-1	100	83	100	100	100	100	100	79	100	100	100	
AG-Bb, AG-Ba, AG-Q	85			77	66		63				57	
AG-1 IA, AG-1 IB, AG-1 IC	91	81		99			71			100		
AG-2-2, AG-2-2 IV	100			100	100		100			100	99	
AG-4 HGIII	100			100	100		100			100	98	
AG-4 HGII	100			100	100		92			94	55	
AG-4 HGI, AG 4-HGII	97			91	97		87			62	60	
AG-4 HGIII, AG-4 HGI, AG-4 HGII	80			88	83		79			89	51	
AG-6 HG-I, AG-6 HGV	91		67	94	99	91	57		67	87	89	91
AG-BI	100	100		100	100		100	100		100	88	
AG-8(ZG1-2), AG-8(ZG1-3), AG-8(ZG1-5)	100			100	82		95			97		
AG-10	100			100	99		100			100	73	
AG-F, CAG-5	100			100	100		100			99	97	

Bootstrap percentages for the subset of the clades supported by bootstrap values above 50% in at least five analyses for each gene partition and in combination of the internal transcribed spacer (ITS), 28S large subunit (LSU), and β-tubulin genes (TUB) when positions with indels are included as ambiguous characters (A), or excluded (B). Empty boxes indicate bootstrap values lower than 50%.

Table 5
Effect of five different indel codings in clade support from combined analyses

CLADES	A	B	C	D	E
AG-D, CAG-1	100	100	100	100	100
AG-A,AG-Bo	100	100	100	100	100
AG-G,AG-L,AG-O	96	92	96	97	98
AG-L,AG-O	59	67	82	79	94
AG-D,CAG-1,AG-Q,AG-Ba,AG-Bb		66	65	63	56
AG-Bb, AG-Ba		52	71	74	73
AG-A, AG-Bo, AG-G, AG-L, AG-O, AG1 to 11 and BI (all *R. solani*), CAG-3, CAG-4, CAG-5, CAG-6, CAG-7, AG-F	80	91	94	87	
AG-A, AG-Bo, AG1 to 11 and BI (all *R. solani*), CAG-3, CAG-4, CAG-5, CAG-6, CAG-7, AG-F		56	53	56	92
AG1 to 11 and BI (all *R. solani*), CAG-3, CAG-4, CAG-5, CAG-6, CAG-7, AG-F	91	75	98	99	100
AG 1-IA,AG 1-IB,AG 1-IC	76		81	82	93
AG 4-HGIII	100	98	100	100	100
AG 4-HGII	100		98	96	100
AG 4-HGI, AG 4-HGII, AG 4-HGIII	83	75	95	92	93
AG-10	99	78	99	99	100
AG 2-2, AG 2-2 IV	100	89	100	100	100
AG 6-HGI, AG 6-HGV	99	97	99	100	100
AG-BI	100	100	100	100	100
AG-F,CAG-5	100	98	100	100	100
AG-8(ZG1-3),AG-8(ZG1-2),AG-8(ZG1-5)	94		70	63	92
AG-8(ZG1-2),AG-8(ZG1-5)		91	77	72	80

Bootstrap percentages for subset of the clades supported by bootstrap values above 50% in at least four analyses with three gene partitions (ITS, LSU, and β-tubulin) and indels with different coding schemes: (A) as ambiguous characters; (B) excluded; (C) multistate for different sequence; (D) multistate for different length; and (E) as different characters for each distinct sequence. Empty boxes indicate bootstrap values lower than 50%.

has been previously suggested by other researchers based on an examination of septal pore, teleomorph characters and sequence data (Andersen, 1996; González et al., 2001; Moore, 1996; Müller et al., 1998; Talbot, 1970; Tu and Kimbrough, 1978). Our analyses also revealed that clades within *R. solani* and binucleate *Rhizoctonia* corresponded with known groups based on hyphal anastomosis. However, levels of support and relationships among groups are sensitive to gene partition and to the treatment of gaps. Among five strategies used for coding the indels, coding them as different characters for each distinct sequence best recovers the phylogenetic information. Bootstrap and Bremer support on basal branches are the highest when this scoring scheme was used (Fig. 3A).

The phylogeny depicted in Fig. 3B reveals that transition to a parasitic stage might have occurred from a saprophytic binucleate ancestor. However, most clades revealed in this study are composed indistinctly of saprophytic and patho-

genic isolates with different plant host. Such heterogeneity indicates neither pathogenicity nor plant hosts are good characters for distinguishing species within the *R. solani* species complex. Isolates of this fungus have been described as pathogenic on over 200 different plant species worldwide with symptoms ranging from damping-off to stem cankers (Anderson, 1982). According to our phylogenetic analyses, it seems that host shifts must have occurred several times in the evolution of *Rhizoctonia* species (*Thanatephorous* and *Ceratobasidium* teleomorphs). Our analyses also revealed that clades are not correlated with geographical origin. Most clades with high bootstrap and Bremer support indices include mixed *Rhizoctonia* isolates from Europe, North America, Australia, or Asia indicating that AG's from this broadly distributed species complex do not have preference for a geographic location. Only clades formed with isolates of AG6 (from Japan) and AG8 (from Australia) had same geographical origin in this analysis (Fig. 3).

Morphological characteristics such as mycelia coloration, hyphal diameter, number of nuclei, length of cells, shape, and size of monilioid cells, and sclerotial size have generally been used in the characterization of *Rhizoctonia* spp. However, these features vary considerably with temperature, light and composition of the medium (Andersen, 1990; Stalpers and Andersen, 1996). Therefore many isolates of *Thanatephorus* and *Ceratobasidium* often have a similar appearance when grown on nutrient medium, and even hyphal fusion among unrelated isolates can occur (Burpee et al., 1980; Kotila, 1929; Ogoshi, 1987; Parmeter et al., 1967). Besides, nuclear number in vegetative cells is rarely absolute. For example, *Thanatephorus* (anamorph = *R. solani*) is considered to be multinucleate (Parmeter and Whitney, 1970), but nuclear condition is variable, cells of a single strain may display a large variation in number (e.g., 2–18) (Tu et al., 1977). Moreover, the hyphal cells of *Ceratobasidium* (anamorph = *Rhizoctonia* spp.) are generally binucleate (Burpee et al., 1980; Parmeter et al., 1967), but *C. koleroga* has been reported to be multinucleate (Stalpers and Andersen, 1996). This evidence and our results with sequence data suggest that some isolates might be better classified within *Thanatephorus* than within *Ceratobasidium* based on nuclear condition and hyphal anastomosis reaction. Transition from binucleate to multinucleate hyphal cells seems to have occurred at least two times independently, as indicated in this study. However, more testing with additional isolates of *Ceratobasidium* is required to substantiate the relationships of the anastomosis groups CAG-3, CAG-4, CAG-5, CAG-6, CAG-7, and AG-F with *Thanatephorus*. Overall, it appears that sequence data supports natural groups within *Rhizoctonia* fungi with *Thanatephorous* and *Ceratobasidium* teleomorphs better than other characters used in the past such as number of nuclei, plant host or morphology.

In general, cladistic analyses of individual and combined data sets from three gene partitions (ITS, LSU rDNA, and β-tubulin) with indels included, excluded or recoded, provide similar conclusions about relationships of *Rhizoctonia* anamorphs of *Ceratobasidium* as have been described previously (Cubeta et al., 1991; González et al., 2001). For example, AG-A and AG-Bo (*C. cornigerum*); AG-Ba (*C. setariae*, anamorph = *R. fumigata*) and AG-Bb (*C. oryzae-sativae*); AG-D and CAG-1 (*C. cereale* Murray and Burpee (anamorph = *R. cerealis* Van der Hoeven)); AG-F and CAG-5 (*Ceratobasidium* sp.); and AG-L and AG-O (*Ceratobasidium* sp.) appear to represent independent evolutionary lineages that correspond to different species of *Ceratobasidium*. The results of this study also provide evidence for the monophyly of six AG's (AG-1, AG-4, AG-6, AG-8, AG-10, and AG-BI) within *R. solani* anamorphs of *Thanatephorus* and for the polyphyletic origin of AG-2. Additionally, results indicate that coding indels as different characters for each distinct sequence (Coding E) best recovers the phylogenetic information. DD scores, bootstrap and Bremer support on basal branches are the highest when this scoring scheme was used. The lowest DD score, bootstrap and Bremer support were observed for the three gene partitions combined where indels were excluded from the analysis (Coding B).

The clades revealed in the analyses when indels are coded as different characters (Coding E) represent at least 17 species within *Thanatephorus* and *Ceratobasidium* teleomorphs. Most of these clades are congruent with current groupings based on hyphal anastomosis behavior. The lack of stability within anastomosis groups in the different analyses performed led to conclude that smaller clades within *R. solani* (at subgroup level) represent genetic variation within anastomosis groups and not taxonomic species.

The proper choice of outgroup for character polarization within the study group was problematic. Ideally, the sister group has to be included as the outgroup. However, when phylogenetic relationships are uncertain, it is possible to incorrectly select an outgroup taxon that may even be derived from within the ingroup (Hopple and Vilgalys, 1999). In the *R. solani* species complex there was not prior knowledge of a putative sister group, therefore preliminary higher and lower-level analyses had to be conducted to provide insight into the appropriate choice of terminal units from among the current data set for use in orienting the trees. Our lower-level phylogenetic analysis revealed that the suitable isolate was a binucleate *Rhizoctonia* sp. (AG-H) (teleomorph = *Ceratobasidium*). However, the same preliminary analysis also revealed that fungi within *Rhizoctonia* are polyphyletic because some isolates of *Uthatobasidium* were mixed and dispersed across several clades within *Thanatephorus* and *Ceratobasidium*. Therefore, it is necessary to do more extensive analyses to better address their relationships.

The results of this study did not fully resolve the problem of relationships and classification of fungi within the *R. solani* species complex. The fact that clades were not stable in different analyses showed that more taxa and perhaps another region of the genome should be examined to obtain a more reliable phylogeny (Graybeal, 1998; Poe, 1998), and to define more precisely how many species are there in these anastomo-

sis groups within the *R. solani* species complex based on their phylogenetic relationships. At present, we cannot reject the hypothesis that each anastomosis group within *R. solani* (teleomorph = *Thanatephorus*) represents a genetically isolated group. However, a conservative approach is only to recommend that four clades corresponding to AG-1, AG-4, AG-6, and AG-8 of *R. solani* (teleomorph = *Thanatephorus*) be recognized as four different phylogenetic species (Fig. 3).

The diversity of sequence data obtained for the *R. solani* species complex has provided an opportunity to evaluate the effect of indels on the stability of phylogenetic hypotheses. The coding approach used in this study may be useful for other data sets involving a complex of closely related species that have differences in sequence length of genes examined and used for phylogenetic inferences. The precise strategy for coding indels for phylogenetic studies still remains to be established, but judging from the results obtained in this study, it seems that coding them as different characters for each distinct sequence (Coding E) recovered historical information. Therefore, cladistic analyses with gene sequences of different length should evaluate the effects of including indels on the stability of phylogenetic hypothesis. Such examination provides an opportunity to test indels as hypotheses of homology to the cladistic test of congruence in parsimony analyses. Considering our results we conclude that although not devoid of homoplasy, indels can be useful markers of shared history at lower taxonomic levels.

Acknowledgments

We thank two anonymous reviewers, David Hibbett and Efraín De Luna for their critical comments on the manuscript. This research was supported in part, by Instituto de Ecología, A.C. account number 902-10-358 to D.G.

References

Aagesen, L., Petersen, G., Seberg, O., 2005. Sequence length variation, indel costs, and congruence in sensitivity analysis. Cladistics 21, 15–30.

Andersen, T.F., 1990. A study of hyphal morphology in the form genus *Rhizoctonia*. Mycotaxon 37, 25–46.

Andersen, T.F., 1996. A comparative taxonomic study of *Rhizoctonia sensu lato* employing morphological, ultrastructural and molecular methods. Mycol. Res. 100, 1117–1128.

Anderson, N.A., 1982. The genetics and pathology of *Rhizoctonia solani*. Annu. Rev. Phytopathol. 20, 329–347.

Burpee, L.L., Sanders, P.L., Junior, H.C., Sherwood, R.T., 1980. Anastomosis groups among isolates of *Ceratobasidium cornigerum* and related fungi. Mycologia 72, 689–701.

Coetzee, M.P.A., Bloomer, P., Ridley, G.S., Wingfield, M.J., 2003. Molecular identification and phylogeny of *Armillaria* isolates from South America and Indo-Malaysia. Mycologia 95, 285–293.

Cubeta, M.A., Echandi, E., Abernethy, T., Vilgalys, R., 1991. Characterization of anastomosis groups of binucleate *Rhizoctonia* species using restriction analysis of an amplified ribosomal RNA gene. Phytopathology 81, 1395–1400.

Cubeta, M.A., Vilgalys, R., 2000. *Rhizoctonia*. In: Lederberg, J. (Ed.), Encyclopedia of Microbiology. Academic Press, San Diego, pp. 109–116.

González, D., 1996. Codificación de las inserciones-deleciones en el análisis filogenético de secuencias génicas. Bol. Soc. Bot. México 59, 115–129.

González, D., Carling, E.D., Kuninaga, S., Vilgalys, R., Cubeta, M.A., 2001. Ribosomal DNA systematics of *Ceratobasidium* and *Thanatephorus* with *Rhizoctonia* anamorphs. Mycologia 93, 1138–1150.

Graybeal, A., 1998. Is it better to add taxa or characters to a difficult phylogenetic problem? Syst. Biol. 47, 9–17.

Grubisha, L.C., Trappe, J.M., Molina, R., Spatafora, J.W., 2002. Biology of the ectomycorrhizal genus *Rhizopogon*. VI. Re-examination of infrageneric relationships inferred from phylogenetic analyses of ITS sequences. Mycologia 94, 607–619.

Hibbett, D.S., Thorn, R.G., 2001. Homobasidiomycetes. In: McLaughlin, D.J., McLaughlin, E.G., Lemke, P.A. (Eds.), The Mycota VII. Systematics and Evolution. Part B. Springer, Berlin, pp. 121–168.

Hibbett, D.S., Fukumasa-Nakai, Y., Tsuneda, A., Donoghue, M.J., 1995. Phylogenetic diversity in shiitake inferred from nuclear ribosomal DNA sequences. Mycologia 87, 618–638.

Higgins, D.G., Bleasby, A.J., Fuchs, R., 1992. CLUSTAL V: improved software for multiple sequence alignment. CABIOS 8, 189–191.

Hopple Jr., J.S., Vilgalys, R., 1999. Phylogenetic relationships in the mushroom genus *Coprinus* and dark-spored allies based on sequence data from the nuclear gene coding for the large ribosomal subunit RNA: divergent domains, outgroups, and monophyly. Mol. Phylogenet. Evol. 13, 1–19.

Kitching, I.L., Forey, P.L., Humphries, C.J., Williams, D.M., 1998. Cladistics: The Theory and Practice of Parsimony Analysis. Oxford University Press, New York.

Kotila, J.E., 1929. A study of biology of a new spore-forming *Rhizoctonia, Corticium praticola*. Phytopathology 19, 1059–1099.

Kropp, B.R., Hansen, D.R., Wolf, P.G., Flint, K.M., Thomson, S.V., 1997. A study on the phylogeny of Dyer's woad rust fungus and other species of *Puccinia* from crucifers. Phytopathology 87, 565–571.

Kuninaga, S., Natsuaki, T., Takeuchi, T., Yokosawa, R., 1997. Sequence variation of the rDNA ITS regions within and between anastomosis groups in *Rhizoctonia solani*. Curr. Genet. 32, 237–243.

Lutzoni, F., Wagner, P., Reeb, V., Zoller, S., 2000. Integrating ambiguously aligned regions of DNA sequences in phylogenetic analyses without violating positional homology. Syst. Biol. 49, 628–651.

Matsumoto, T., Yamamoto, W., Hirane, S., 1932. Physiology and parasitology of the fungi generally referred to as *Hypochnus sasakii* Shirai. I. Differentiaton of the strains by mean of hyphal fusion and culture in differential media. J. Soc. Trop. Agric. 4, 370–388.

Milinkovitch, M.C., Meyer, A., Powell, J.R., 1994. Phylogeny of all major groups of cetaceans based on DNA sequences from three mitochondrial genes. Mol. Biol. Evol. 11, 939–948.

Moncalvo, J.-M., Vilgalys, R., Redhead, S.A., Johnson, J.E., James, T.Y., Aime, M.C., Hofstetter, V., Verduin, S.J.W., Larsson, E., Baroni, T.J., Thorn, R.G., Jacobsson, S., Clémençon, H., Miller Jr., O.K., 2002. One hundred and seventeen clades of euagarics. Mol. Phylogenet. Evol. 23, 357–400.

Moore, R.T., 1996. The dolipore/parenthesome septum in modern taxonomy. In: Sneh, B., Jabaji-Hare, S., Neate, S., Dijst, G. (Eds.), *Rhizoctonia* Species: Taxonomy, Molecular Biology, Ecology, Pathology and Disease Control. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 13–35.

Müller, W.H., Stalpers, J.A., van Aelst, A.C., vanderKrift, T.P., Boekhout, T., 1998. Field emission gun-scanning electron microscopy of septal pore caps of selected species in the *Rhizoctonia* s.l. complex. Mycologia 90, 170–179.

Ogoshi, A., 1987. Ecology and pathogenicity of anastomosis and intraspecific groups of *Rhizoctonia solani* Kühn. Annu. Rev. Phytopathol. 25, 125–143.

Olsen, G.J., 1988. Phylogenetic analysis using ribosomal RNA. Methods Enzymol. 164, 793–838.

Olsen, G.J., Woese, C.R., 1993. Ribosomal RNA: a key to phylogeny. FASEB J. 7, 113–123.

Parmeter Jr., J.R., Whitney, H.S., Platt, W.D., 1967. Affinities of some *Rhizoctonia species* that resemble mycelium of *Thanatephorus cucumeris*. Phytopathology 57, 218–223.

Parmeter Jr., J.R., Whitney, H.S., 1970. Taxonomy and nomenclature of the imperfect state. In: Parmeter, J.R. (Ed.), Rhizoctonia solani: Biology and Pathology Univ of Calif Press, Berkeley, pp. 7–19.

Poe, S., 1998. Sensitivity of phylogeny estimation to taxonomic sampling. Syst. Biol. 47, 18–31.

Rehner, S.A., Buckley, E., 2005. A *Beauveria* phylogeny inferred from nuclear ITS and EF1-α sequences: evidence for cryptic diversification and links to *Cordyceps* teleomorphs. Mycologia 97, 84–98.

Sanderson, M.J., 1989. Confidence limits on phylogenies: the bootstrap revisited. Cladistics 5, 113–129.

Schultz, H., 1936. Vergleichende Untersuchungen zur Okologic, Morphologie und Systematik des Vermchrungspilzes. Abr.Biol. Reichsanst. Land-u Forstwirtsch Berl 22, 1–41.

Simmons, M.P., Ochoterena, H., 2000. Gaps as characters in sequence-based phylogenetic analyses. Syst. Biol. 49, 369–381.

Soltis, D., Soltis, P.S., Chase, M.W., Mort, M.E., Albach, D.C., Zanis, M., Savolainen, V., Hahn, W.H., Hoot, S.B., Fay, M.F., Axtell, M., Swensen, S.M., Prince, L.M., Kress, W.J., Nixon, K.C., Farris, J.S., 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbc*L, and *atp*B sequences. Bot. J. Linn. Soc. 133, 381–461.

Stalpers, J.A., Andersen, T.F., 1996. A synopsis of the taxonomy of teleomorphs connected with *Rhizoctonia* S.L. In: Sneth, B., Jabaji-Hare, S., Neate, S., Dijst, G. (Eds.), *Rhizoctonia* Species: Taxonomy, Molecular Biology, Ecology, Pathology and Disease Control. Kluwer Academic Publishers, Dordrecht, pp. 49–63.

Swofford, D.L., 2000. Phylogenetic analysis using parsimony (* and other methods). 4.0b4. Sunderland, Massachusetts. Sinauer Associates.

Talbot, P.H.B., 1970. Taxonomy and nomenclature of the perfect state. In: Parmeter, J.R. (Ed.), *Rhizoctonia solani*: Biology and Pathology. Univ. of Calif. Press, Berkeley.

Tu, C.C., Kimbrough, J.W., Aldrich, H.C., 1977. Cytology and ultrastructure of *Thanatephorus cucumeris* and related taxa of the *Rhizoctonia* complex. Can. J. Bot. 5, 2419–2436.

Tu, C.C., Kimbrough, J.W., 1978. Systematics and phylogeny of fungi in the *Rhizoctonia* complex. Bot. Gaz. 139, 454–466.

Vogler, A.P., DeSalle, R., 1994. Evolution and phylogenetic information content of the ITS-1 region in the tiger beetle *Cicindela dorsalis*. Mol. Biol. Evol. 11, 393–405.